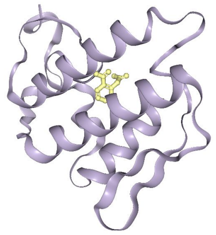
Using Deep Learning with Spatial Transformations to Predict Protein-Ligand Binding

David Ban

North Allegheny Senior High School

Introduction - Drug Discovery

- Time Consuming - Expensive - Labor Intensive
- Protein-Ligand Binding is Important for Drug Discovery
    - 90% of drugs on market act on proteins*
- Protein: molecules essential for body functions
- Ligand: molecule that binds to a protein, i.e. drug



1OXR - Aspirin binding

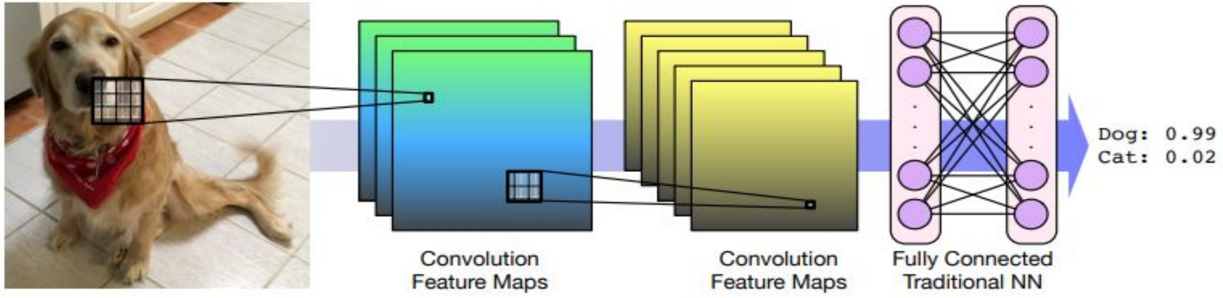
Causes anti-inflammatory effects

Singh RK, Ethayathulla AS, Jabeen T, Sharma S, Kaur P, Singh TP.J Drug Target. 2005 Feb; 13(2):113-9.

*Rask-Andersen, et al., Trends in the exploitation of novel drug targets. Nature Reviews Drug Discovery, 10:579-590, Aug. 2011

Introduction - Convolutional Neural Networks

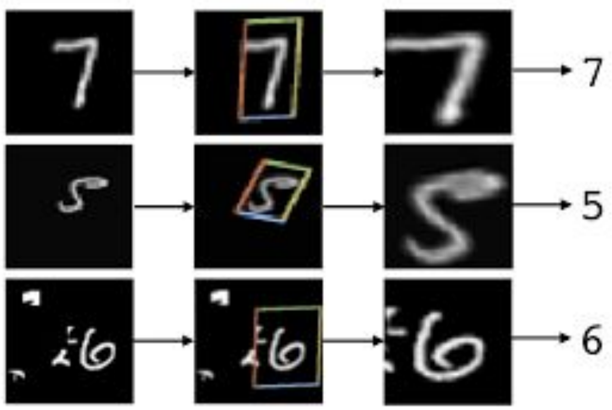
- Deep Learning Neural networks
    - Commonly applied to analyzing visual imagery
    - Learns features of the image as “filters” or “feature maps"
    - Each additional layer allows learning at a higher level features



Ragoza, M. et al. Protein-Ligand Scoring with Convolutional Neural Networks. J. Chem. Information and Modeling. 942-957, April. 2017

Introduction - Spatial Transformer Networks

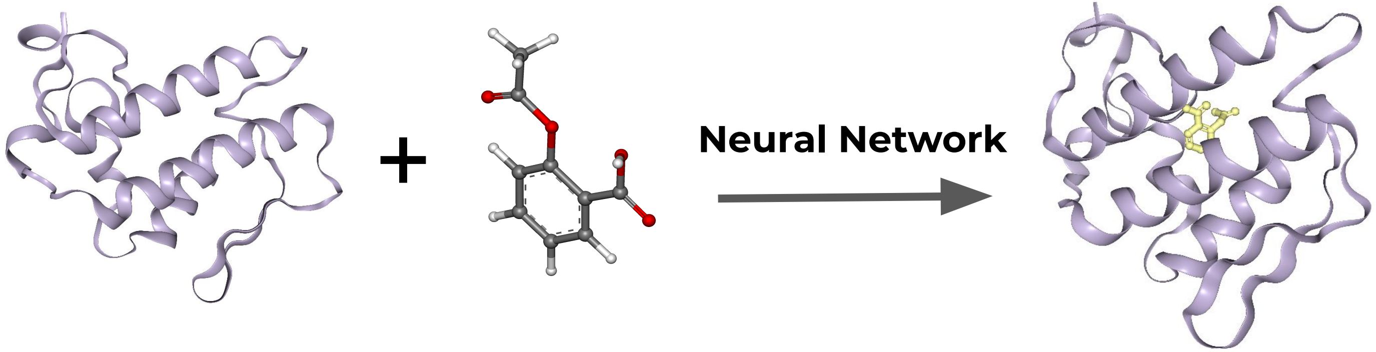
- STNs module incorporated into CNN
- Manipulates data in multiple dimensions
- Spatially transforms data
- Resilient to perturbation
  - translation, rotation, dilation, warping



Jaderberg, M. et al. “Spatial Transformer Networks”. Feb. 2016

Introduction - Can we do it?

Question: can neural networks predict protein-ligand binding after rigid body transformations?

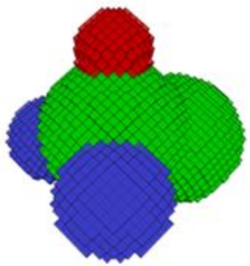


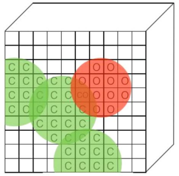
1OXR Aspirin 1OXR-Aspirin

Method - Ligand Voxelization and Perturbation

3D Voxelization

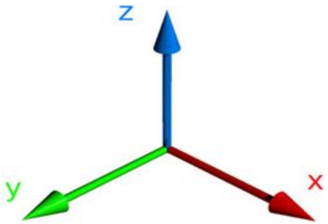
- Represents structures in cartesian coordinates
- Atoms represented as voxels





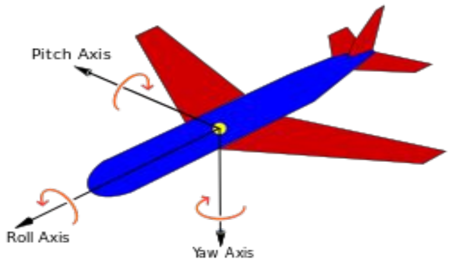
Translation

- Translated along x, y and z axis
- Error measured in angstroms

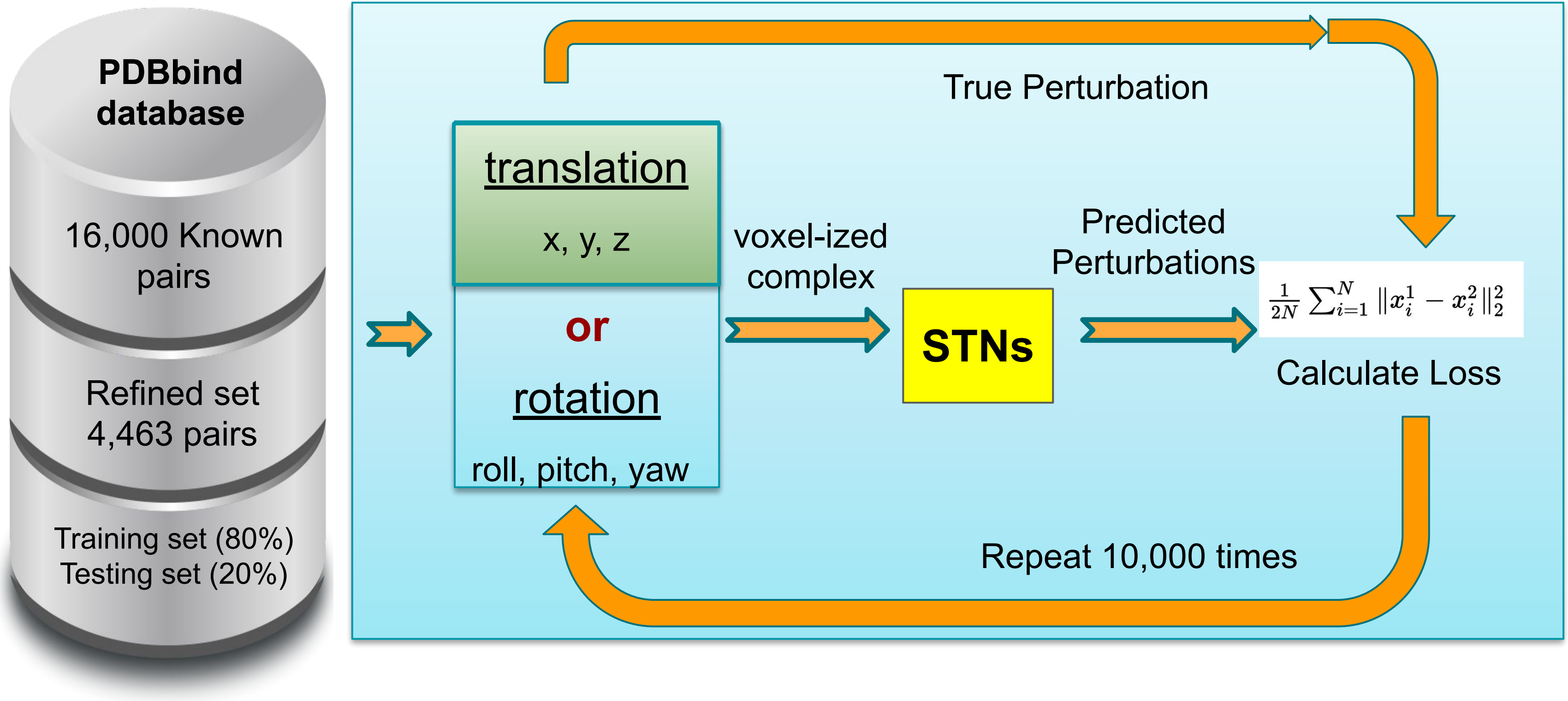


Rotation

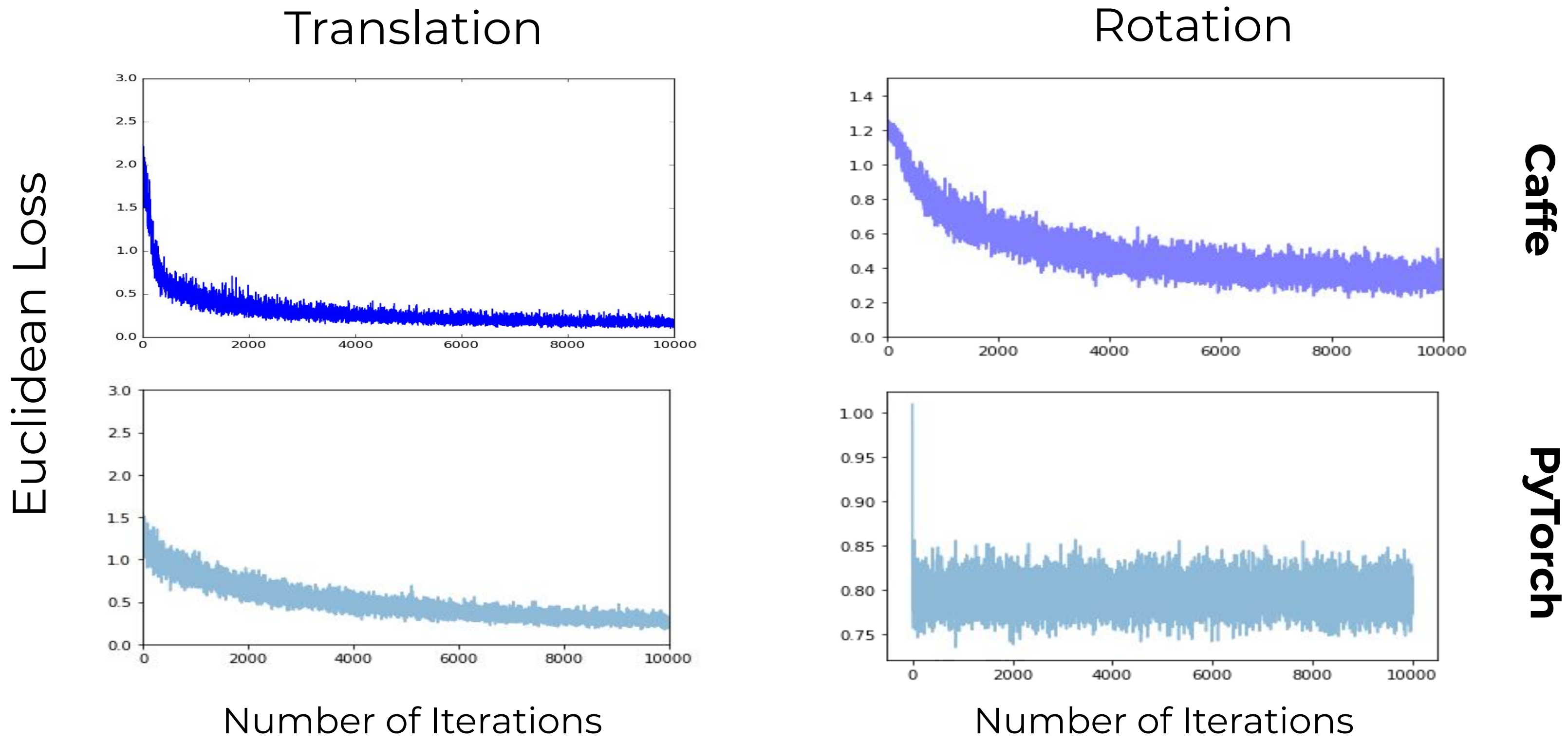
- Translated along roll, pitch and yaw
- Error measured in radians



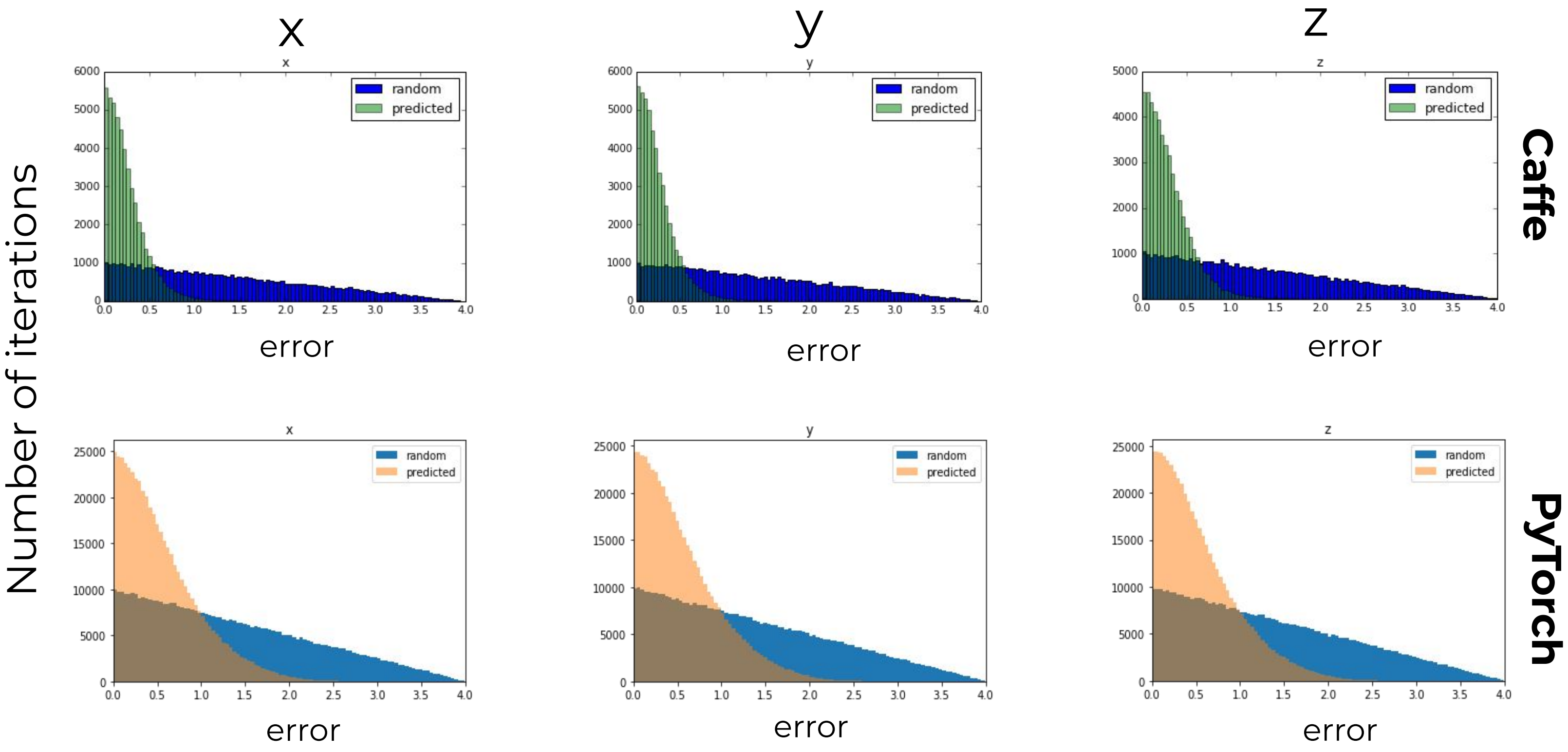
Methods - Pytorch and Caffe Outline



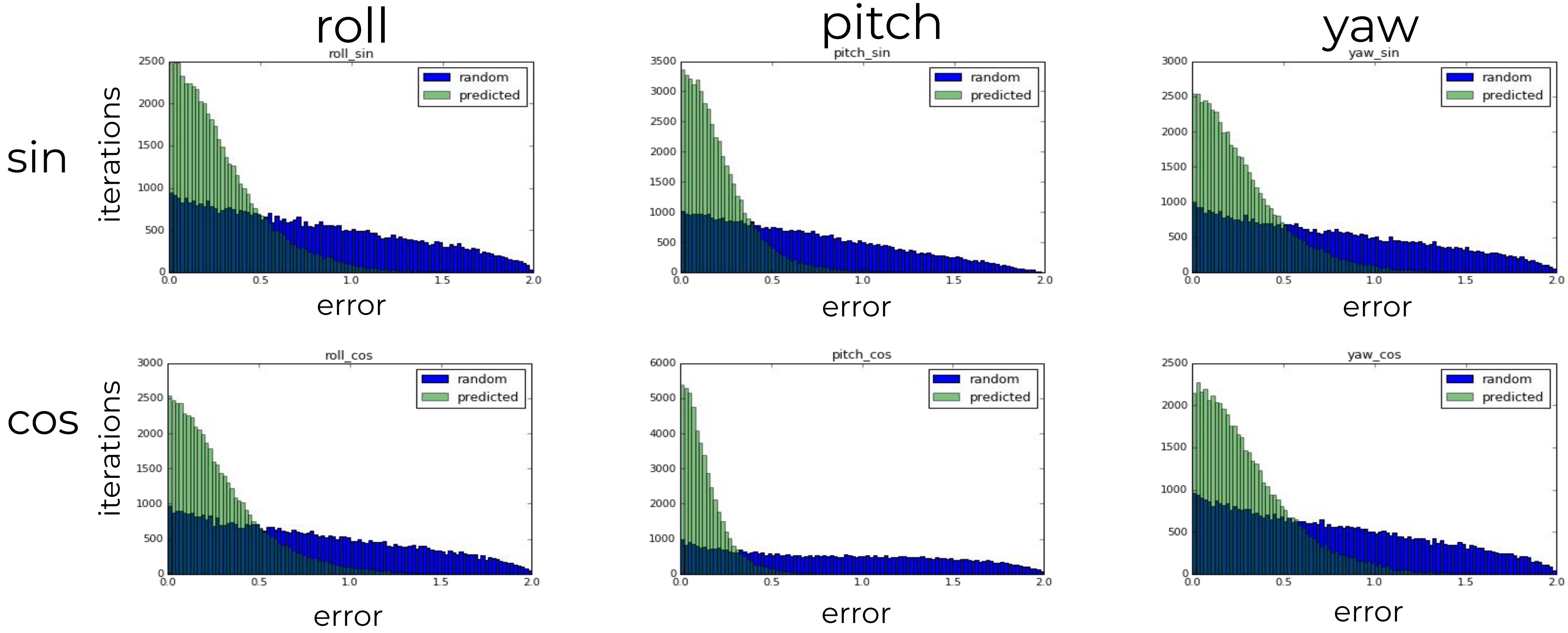
Results - Loss Convergence: Caffe vs PyTorch



Results - Prediction Accuracy: Translation



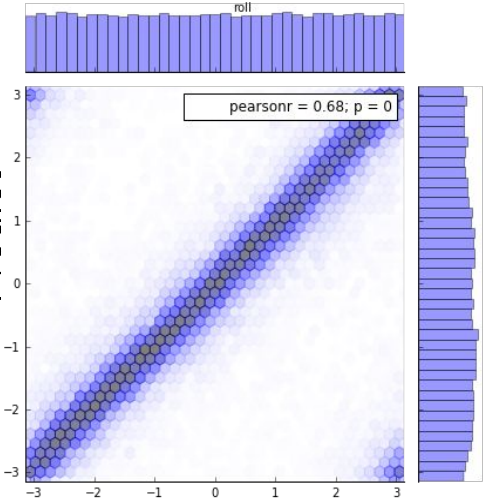
Results - Prediction Accuracy: Rotation-Caffe



Results – Correlation between predicted and true values for rotational models: Caffe

roll

Predict

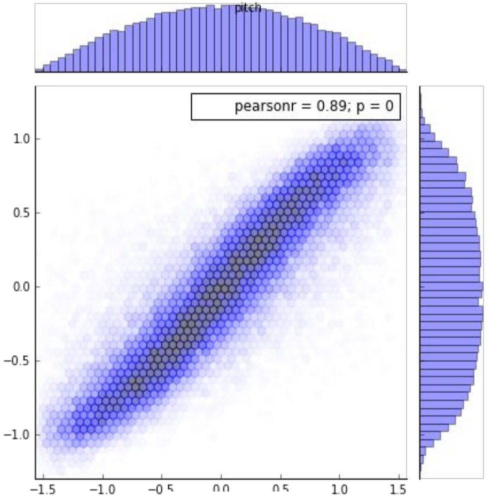


True

Pearson r = 0.68, p = 0

pitch

Predict

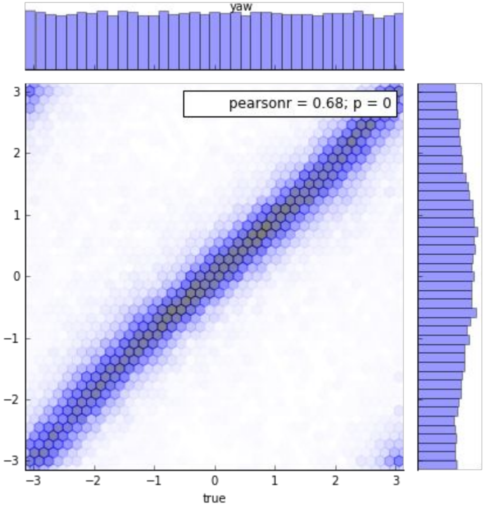


True

Pearson r = 0.89, p = 0

yaw

Predict



True

Pearson r = 0.68, p = 0

Results - Summary of Pearson R

		Roll		Pitch		Yaw		X		Y		Z	
		Train	**Test**	**Train**	**Test**	**Train**	**Test**	**Train**	**Test**	**Train**	**Test**	**Train**	**Test**
Caffe	**2 Angstrom Translation**	-	-	-	-	-	-	0.96	0.96	0.96	0.96	0.96	0.96
	Rotation	0.68*	0.68*	0.89	0.89	0.69*	0.68*	-	-	-	-	-	-
PyTorch	**2 Angstrom Translation**	-	-	-	-	-	-	0.86	0.86	0.85	0.85	0.86	0.86
	Rotation	0	0	0	0	0	0	-	-	-	-	-	-

*Indicates biased Pearson R value

Conclusions

- STNs are able to converge indicating their learning capacity. Caffe model seems to work better than PyTorch model
- STNs are able to learn translational perturbations using Caffe or PyTorch
- Random rotations of the ligand can be predicted by STNs but with weaker accuracy than translation using Caffe
- PyTorch model was not able to converge for rotation perturbations.
- Results are promising - further testing needs to be performed

Acknowledgements

AMIA
INFORMATICS PROFESSIONALS. LEADING THE WAY.

Dr. David R. Koes for his guidance and support throughout my project.

Department of Computational and Systems Biology at the University of Pittsburgh for allowing me to work in Dr. Koes' lab and have access to their supercomputer facilities.

Department of Computational and Systems Biology

Thank you!